Aprendizaje Basado en Instancias

Guillermo Henrión

Temario

- Motivación
- K-Nearest Neighbor
- Locally weighted regression
- Case-based reasoning

Motivación

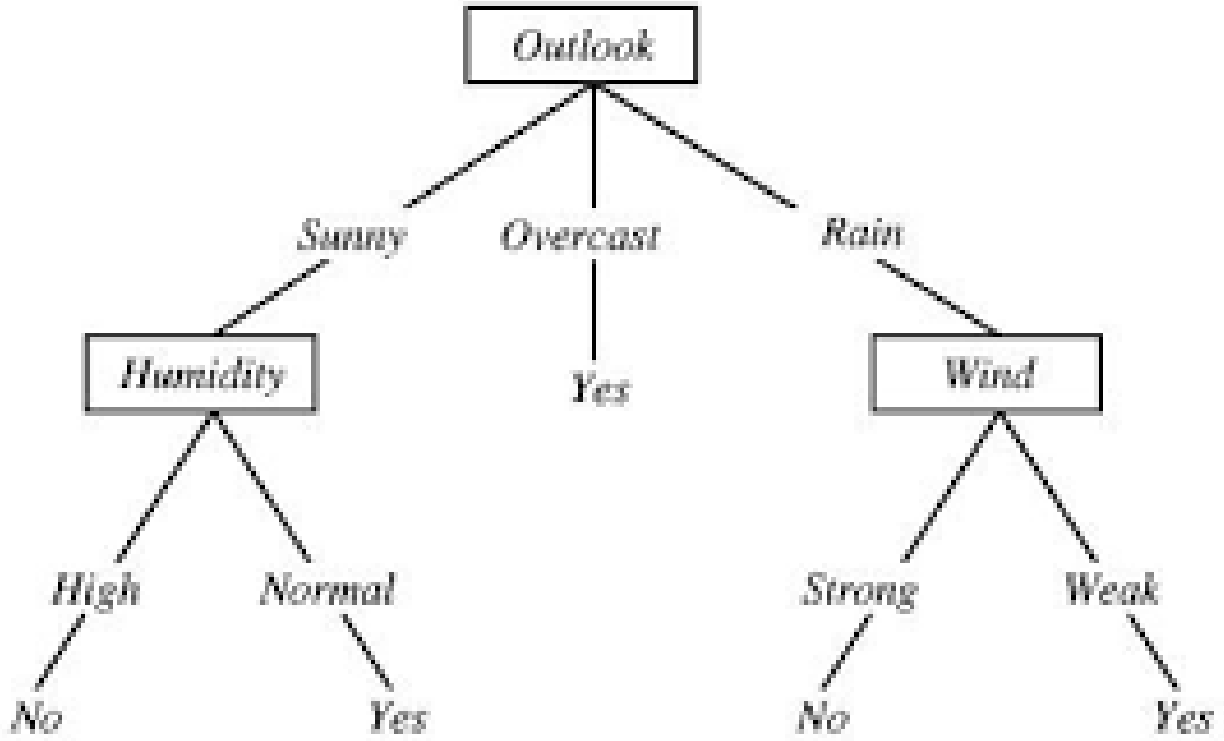
Formas de aprendizaje

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Motivación

Reglas

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No



Motivación

Probabilidades

Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Data	
Temperature	**Play Tennis**
Hot	No
Hot	No
Hot	Yes
Mild	Yes
Cool	Yes
Cool	No
Cool	Yes
Mild	No
Cool	Yes
Mild	Yes
Mild	Yes
Mild	Yes
Hot	Yes
Mild	No

Probability Table			
Temperature	**Play Tennis : Yes**	**Play Tennis : No**	**Probability**
Hot	2	2	4/14 = 0.29
Cool	3	1	4/14 = 0.29
Mild	4	2	6/14 = 0.43
All	9	5	
Probability	9/14 = 0.64	5/14 = 0.36	

Motivación

Cercanía

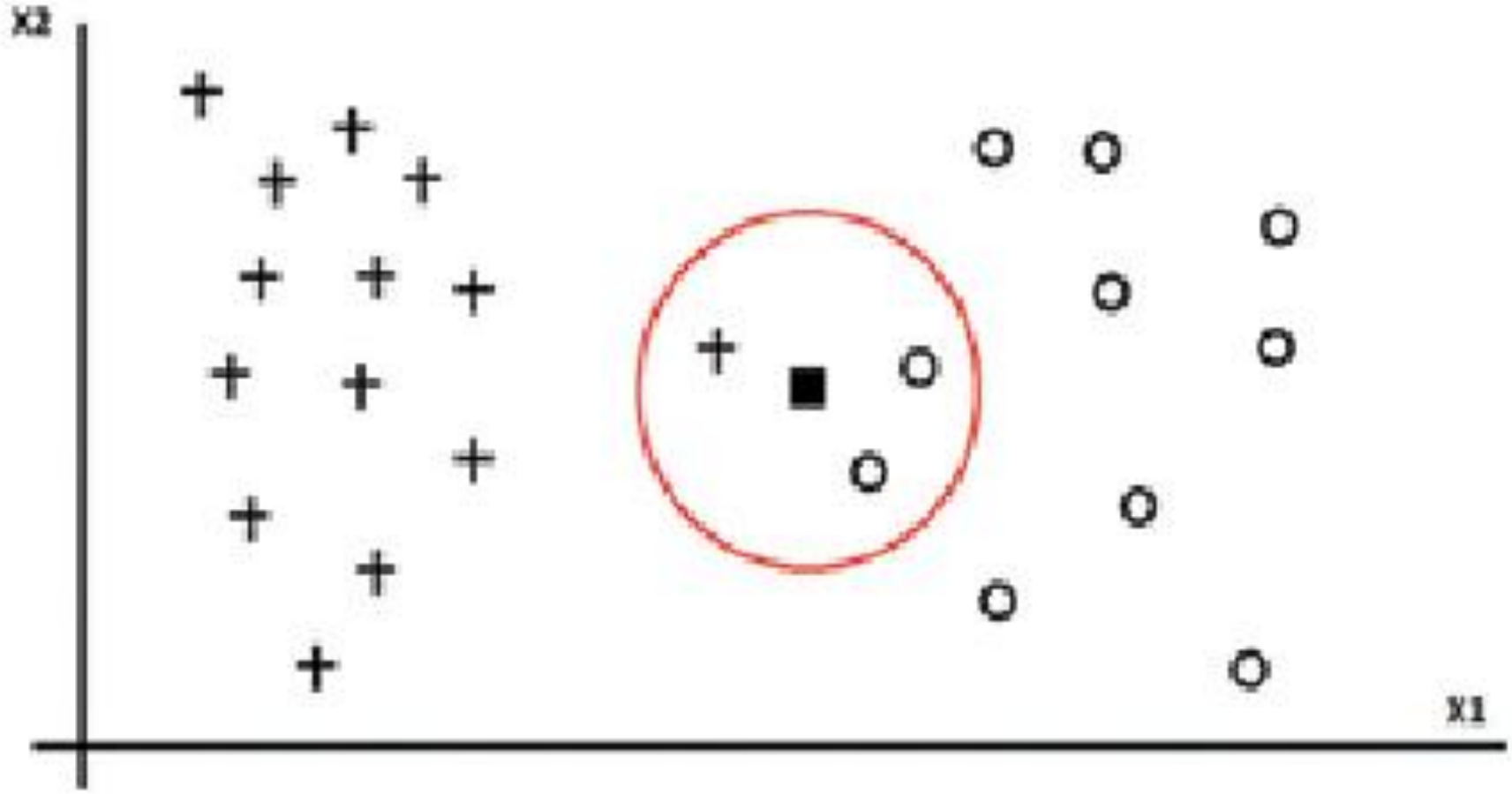
Day	Outlook	Temperature	Humidity	Wind	Play Tennis
1	Sunny	Hot	High	Weak	No
2	Sunny	Hot	High	Strong	No
3	Overcast	Hot	High	Weak	Yes
4	Rain	Mild	High	Weak	Yes
5	Rain	Cool	Normal	Weak	Yes
6	Rain	Cool	Normal	Strong	No
7	Overcast	Cool	Normal	Strong	Yes
8	Sunny	Mild	High	Weak	No
9	Sunny	Cool	Normal	Weak	Yes
10	Rain	Mild	Normal	Weak	Yes
11	Sunny	Mild	Normal	Strong	Yes
12	Overcast	Mild	High	Strong	Yes
13	Overcast	Hot	Normal	Weak	Yes
14	Rain	Mild	High	Strong	No

Entrenamiento

Overcast Mild High Weak Yes

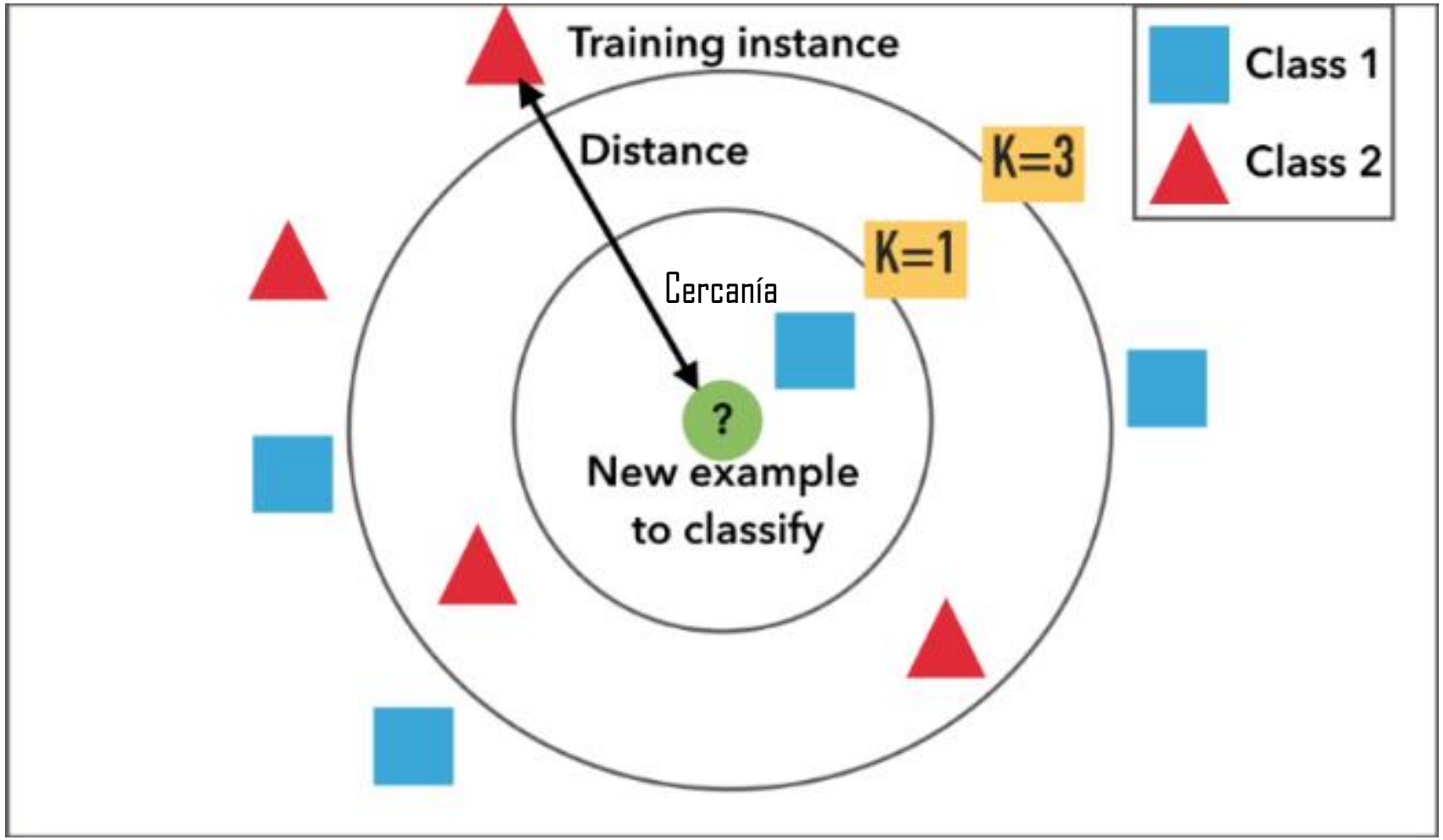
Motivación

Cercanía



Motivación

Cercanía



K-NN

- Es el más simple de los métodos basados en instancias
- Supervisado
- Asume que todas las instancias están en un espacio n-dimensional
- Difiere el proceso de clasificar hasta que una nueva instancia deba ser clasificada
- Puede aprender funciones complejas
- No pierde información
- Puede ser “engañado” por atributos irrelevantes

K-NN

Dada una instancia x y a_i sus atributos

$$\langle a_1(x), a_2(x), \ldots a_n(x) \rangle$$

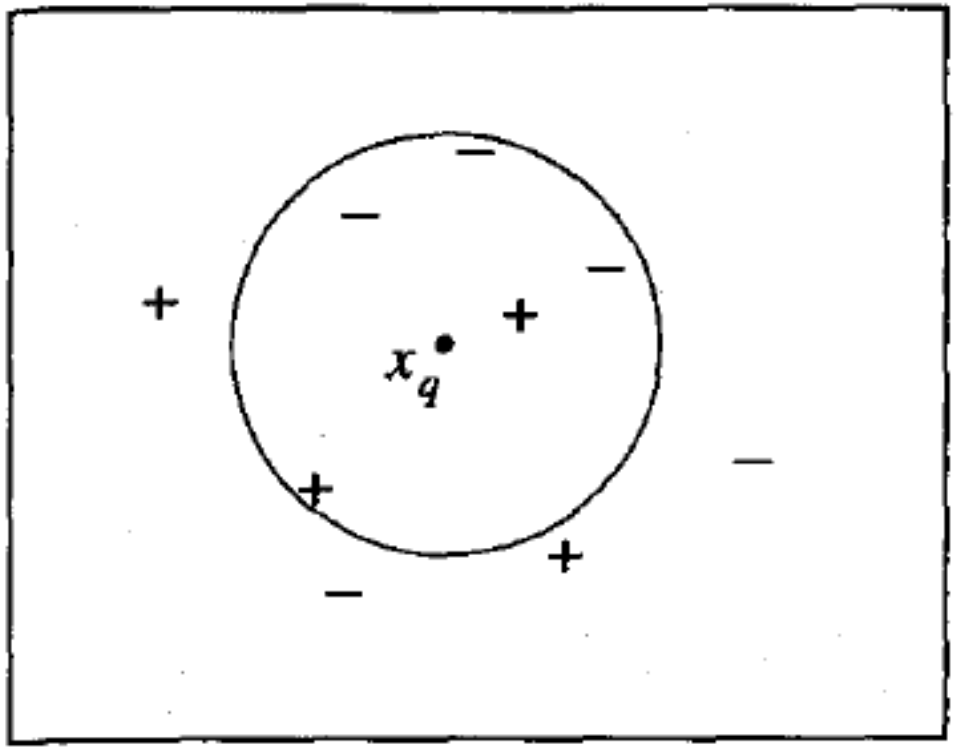
$$d(x_i, x_j) \equiv \sqrt{\sum_{r=1}^{n} (a_r(x_i) - a_r(x_j))^2}$$

K-NN

funciones discretas

$$f : \Re^n \rightarrow V \{v_1, \ldots v_s\}$$

Luego el K-nn retorna para *f*(xq) el valor más común de *f* entre los k ejemplos más cercanos a xq



K-NN

Training algorithm:

- For each training example $\langle x, f(x) \rangle$, add the example to the list *training_examples*

Classification algorithm:

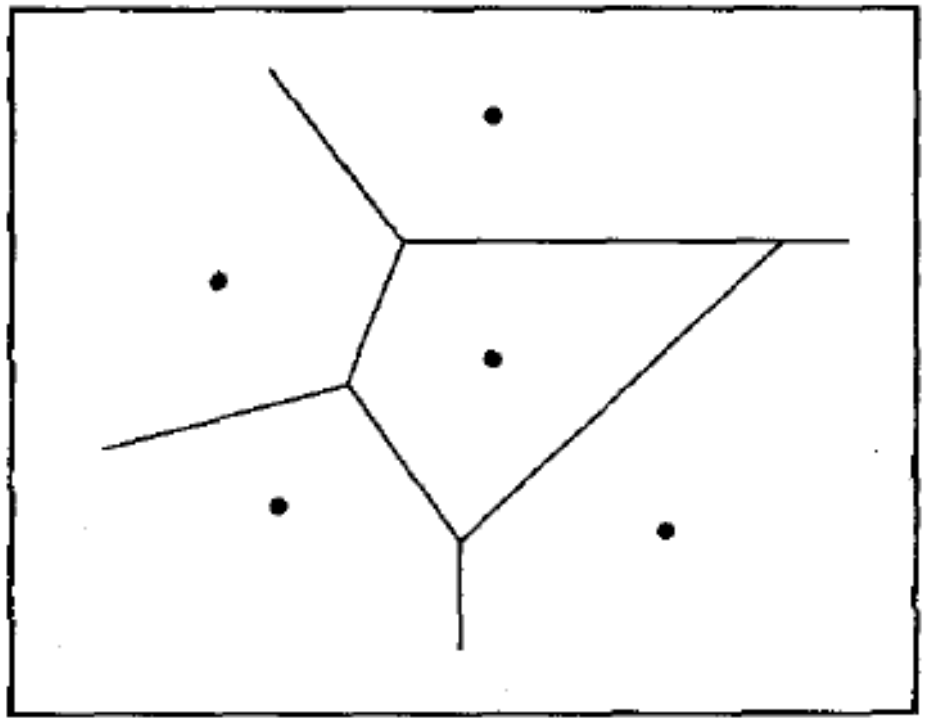
- Given a query instance x_q to be classified,
  - Let $x_1 \ldots x_k$ denote the k instances from *training_examples* that are nearest to x_q
  - Return

$$\hat{f}(x_q) \leftarrow \underset{v \in V}{\operatorname{argmax}} \sum_{i=1}^{k} \delta(v, f(x_i))$$

where $\delta(a, b) = 1$ if $a = b$ and where $\delta(a, b) = 0$ otherwise.

K-NN

Diagrama de Voronoi



La superficie de decisión es una combinación poliedros convexos rodeando cada uno de los ejemplos de entrenamiento

K-NN

funciones continuas

$$f : \Re^n \rightarrow \Re$$

$$\hat{f}(x_q) \leftarrow \frac{\sum_{i=1}^{k} f(x_i)}{k}$$

K-NN

Distance-weigthed

Se pesa el voto de cada vecino de acuerdo a la inversa de su distancia a la instancia a clasificar

$$w_i \equiv \frac{1}{d(x_q, x_i)^2}$$

Discreto

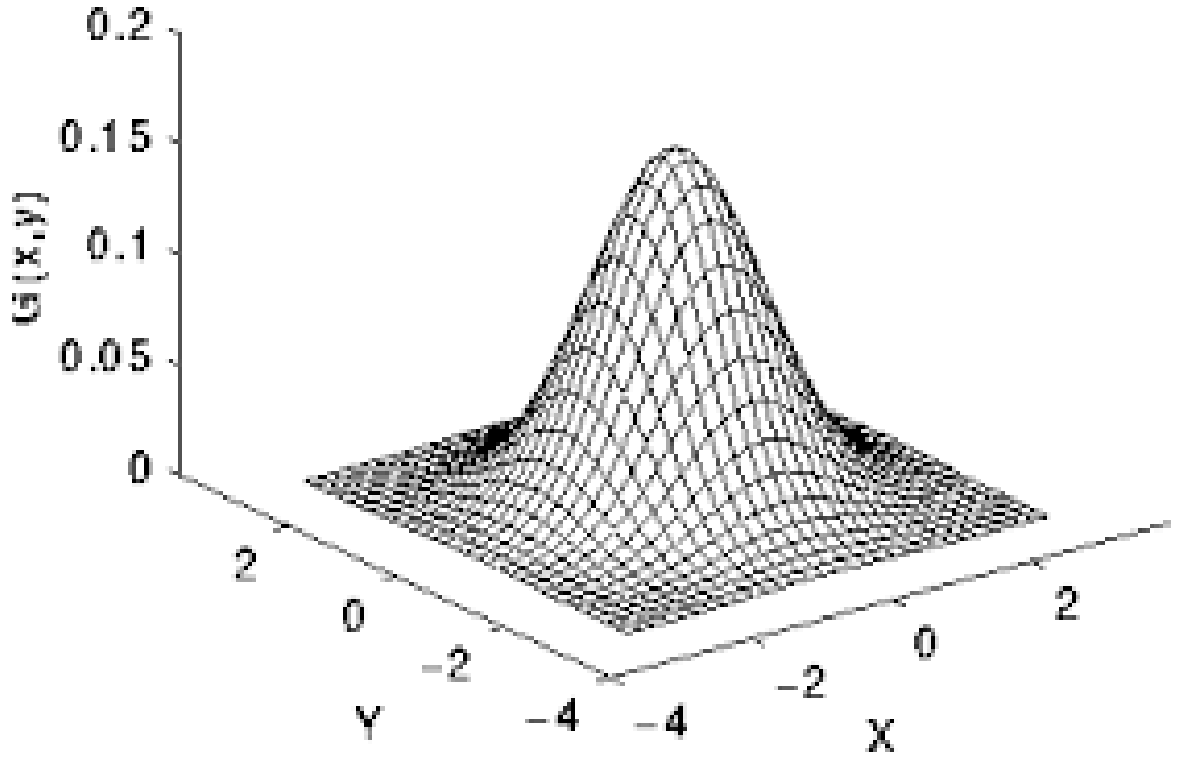
$$\hat{f}(x_q) \leftarrow \underset{v \in V}{\operatorname{argmax}} \sum_{i=1}^{k} w_i \delta(v, f(x_i))$$

Continuo

$$\hat{f}(x_q) \leftarrow \frac{\sum_{i=1}^{k} w_i f(x_i)}{\sum_{i=1}^{k} w_i}$$

K-NN

Distance-weigthed



Locally weighted regression

Knn forma una aproximación local para cada x_q.

Podemos formar una aproximación explícita $f(x)$ para la región alrededor de x_q. ajustando una función (lineal, cuadrática u otra) a sus k vecinos cercanos, pesada por su distancia, y evaluar la nueva instancia x_q con esa función local.

Locally weighted regression

$$\hat{f}(x) = w_0 + w_1 a_1(x) + \cdots + w_n a_n(x)$$

$$E \equiv \frac{1}{2} \sum_{x \in D} (f(x) - \hat{f}(x))^2$$

Locally weighted regression

$$E_1(x_q) \equiv \frac{1}{2} \sum_{x \in \text{ k nearest nbrs of } x_q} (f(x) - \hat{f}(x))^2$$

$$E_2(x_q) \equiv \frac{1}{2} \sum_{x \in D} (f(x) - \hat{f}(x))^2 \; K(d(x_q, x))$$

$$E_3(x_q) \equiv \frac{1}{2} \sum_{x \in \text{ k nearest nbrs of } x_q} (f(x) - \hat{f}(x))^2 \; K(d(x_q, x))$$

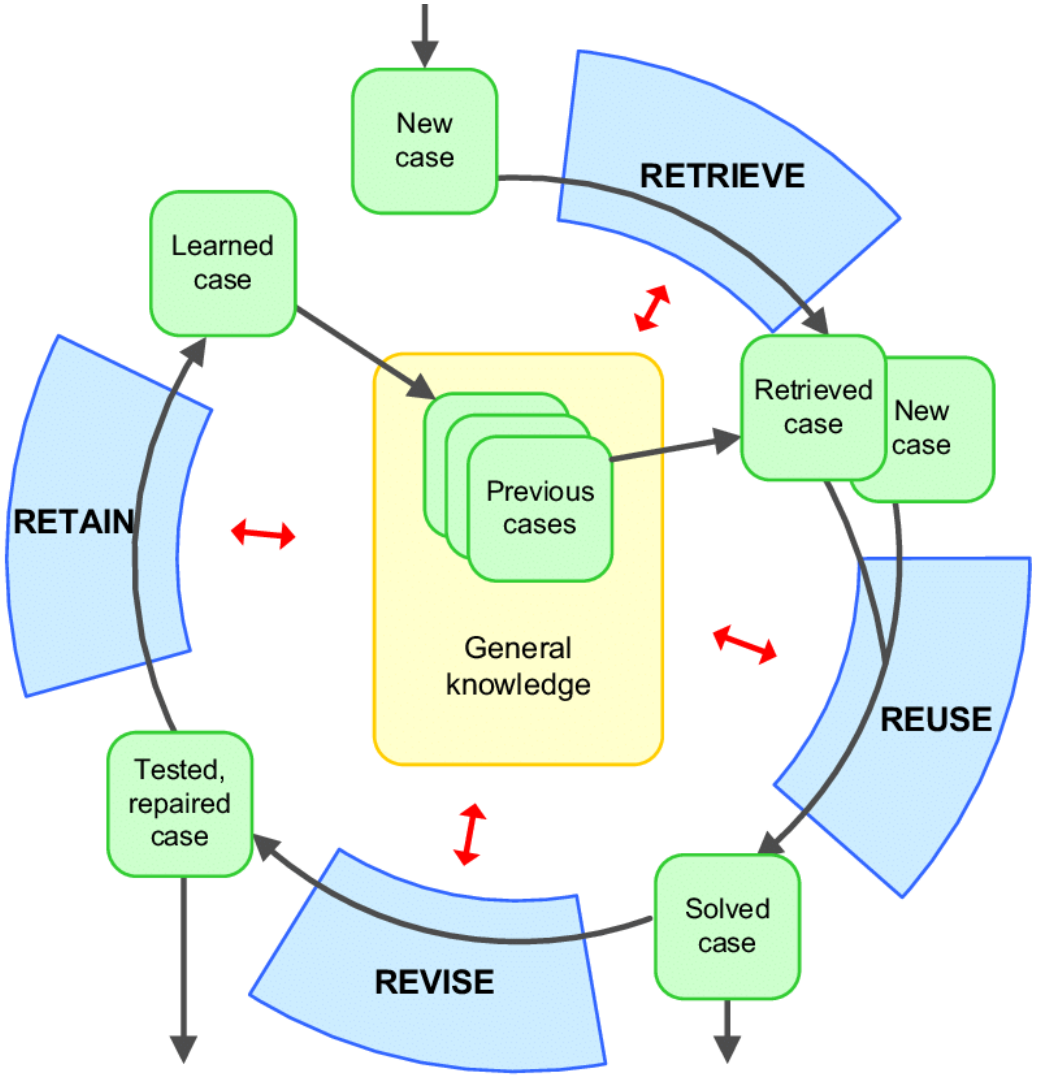
Case-based reasoning

Usa una representación más rica y por ende medidas de distancia más elaboradas

Por ejemplo clasificando textos, debería comparar la similitud de los textos para clasificar un nuevo texto (mediante similitud semántica que usan el significado de las palabras).

O un sistema de toma de decisiones dando respuestas de acuerdo a la similitud con la preguntas. Resuelve problemas basándose en la solución de problemas ya vistos. En este último ejemplo la complejidad pasa por determinar cuándo dos preguntas son parecidas.

Case-based reasoning



Bibliografía

Machine Learning cap. 8 - T. Mitchell

https://web.archive.org/web/20080312053714/http://www.iiia.csic.es/People/enric/AICom.html

https://towardsdatascience.com/how-to-rank-text-content-by-semantic-similarity-4d2419a84c32